

Mycobacterium tuberculosis Protein Tyrosine Phosphatase PtpB Structure Reveals a Diverged Fold and a Buried Active Site

Christoph Grundner, Ho-Leung Ng, and Tom Alber*
Department of Molecular and Cell Biology
University of California, Berkeley
Berkeley, California 94720

Summary

Intracellular pathogenic bacteria manipulate host signal transduction pathways to facilitate infection. *Mycobacterium tuberculosis* protein tyrosine phosphatases (PTPs) PtpA and PtpB are thought to be secreted into host cells and interfere with unidentified signals. To illuminate the mechanisms of regulation and substrate recognition, we determined the 1.7 Å resolution crystal structure of PtpB in complex with the product phosphate. The protein adopts a simplified PTP fold, which combines features of the conventional PTPs and dual-specificity phosphatases. PtpB shows two unusual elaborations—a disordered, acidic loop and a flexible, two-helix lid that covers the active site—that are specific to mycobacterial orthologs. Biochemical studies suggest that substrate mimicry in the lid may protect the phosphatase from oxidative inactivation. The insertion and deletion of large structural elements in PtpB suggest that, outside the active site module, the PTP family is under unusual selective pressure that promotes changes in overall structure.

Introduction

During coevolution with the host immune system, bacterial pathogens have acquired defense mechanisms against host immune effectors. Interfering with host signal transduction is a particularly effective way of manipulating the host cell, because small modifications of signaling events can have large physiologic effects. A number of bacterial signaling proteins have been identified to act in the host. In *Mycobacterium tuberculosis* (*Mtb*), secreted phosphatidyl-inositol-3-phosphate phosphatase inhibits phagosomal maturation (Vergne et al., 2005). The serine/threonine protein kinase PknG from *Mycobacterium bovis* BCG also has been reported to prevent fusion of phagosomes with lysosomes and protect the intraphagosomal bacteria from acidification (Walburger et al., 2004). *Salmonella* secretes the SopE effectors that cause extensive cytoskeletal rearrangements in the host (Fu and Galan, 1998; Kaniga et al., 1996). Once inside the cell, the tyrosine phosphatase SptP reverses the effects of the SopE proteins through its N-terminal GTPase-activating protein segment and restores the normal cytoskeleton structure (Fu and Galan, 1999). *Yersinia* encodes a number of virulence factors (*Yersinia* outer membrane proteins, Yops) that are secreted into the host. YopH, YopE, YopT, and YpkA/YopO interfere with phagocytosis by disrupting the host cytoskeleton, preventing phagocytic uptake

by macrophages (Cornelis, 2002). The YopH tyrosine phosphatase contributes to cytoskeletal disassembly by dephosphorylating focal adhesion complexes (Black and Bliska, 1997). YopH also suppresses the activation of T and B lymphocytes by interfering with early phosphorylation of the antigen receptor complex (Yao et al., 1999), and SptP antagonizes SAPK/JNK activation that might suppress the production of proinflammatory cytokines (Fu and Galan, 1999).

The genome of *Mtb* codes for two tyrosine phosphatases, PtpA and PtpB, but no tyrosine kinase has been found (Cole et al., 1998). In the absence of an endogenous tyrosine kinase, these phosphatases likely target components of host signaling pathways. Several lines of evidence support a function for the *Mtb* tyrosine phosphatase PtpB in the host cell. Despite the absence of any known export signal, PtpB is secreted into the culture medium (Koul et al., 2000). The pH optimum of 5.5 would allow PtpB to function in low-pH compartments of macrophages (Koul et al., 2000). Genetic inactivation of PtpB in *Mtb* caused accelerated mycobacterial cell death in activated macrophages and guinea pigs (Singh et al., 2003). Guinea pigs infected with PtpB-deficient *Mtb* showed an ~70-fold lower bacterial load in the spleen 6 weeks after infection compared to a wild-type infection. This late replication defect coincides with the development of the adaptive immune response, suggesting an involvement of PtpB in evasion of B and T lymphocyte-mediated responses.

Both PtpA and PtpB exhibit protein tyrosine phosphatase activity (Koul et al., 2000). They contain the defining consensus sequences of all protein tyrosine phosphatases, the active site P loop motif, CysX₅Arg, and an essential Asp that functions as a general acid in the catalytic reaction (Barford et al., 1998). While PtpA is classified as a low-molecular weight (LMW) phosphatase, the PtpB sequence falls into the conventional PTP or dual-specificity phosphatase (DSP) class. The conventional PTPs and DSPs contain the P loop in the latter half of the sequence. A histidine immediately preceding the active site cysteine decreases the pKa of the thiol group and makes it a better nucleophile. The catalytic domain consists of a central, twisted β sheet surrounded by α helices. All PTP structures determined to date feature at least a four-stranded parallel β sheet, with up to five additional antiparallel strands flanking the core sheet of most conventional PTPs. In contrast to the DSPs, which dephosphorylate pSer, pThr, and pTyr residues, conventional PTPs act only on pTyr. This narrowed specificity is associated with a larger β sheet (typically up to nine strands) and a deeper active site cleft. Despite little sequence homology, there is a strong structural conservation of the catalytic domain of PTPs. Two representative conventional PTPs, PTP1B and YopH, share only about 20% sequence homology, but high structural similarity (Barford et al., 1994; Stuckey et al., 1994).

For most PTPs, the physiologic substrates are unknown. The human genome codes for 81 functional Tyr kinases and 85 functional PTPs, suggesting specialized

*Correspondence: tom@ucxray.berkeley.edu

Table 1. Data Collection, Phasing, and Refinement Statistics for PtpB

Data Collection and Phasing				
Crystal	Native	Selenomethionine Labeled		
Crystal symmetry	I_{41}	I_{41}		
Unit cell	$a = 113.07 \text{ \AA}, b = 113.07 \text{ \AA}, c = 53.29 \text{ \AA}$	$a = 113.20 \text{ \AA}, b = 113.20 \text{ \AA}, c = 53.70 \text{ \AA}$		
		Peak	Edge	Remote
Wavelength (Å)	1.11587	0.97949	0.97973	0.93927
Resolution (Å)	79–1.7	79–2	79–2	79–2
Completeness (%) ^a	91.2 (48.6)	94.8 (70.7)	95 (72)	93.5 (65.6)
Multiplicity	8.2 (4.5)	4.7 (3.2)	4.7 (3.3)	4.7 (3.4)
R_{merge} (%) ^b	4.1 (35.6)	6.3 (35.1)	5.5 (34.5)	6.1 (42)
$\langle I/\sigma \rangle$ ^a	66.7 (2.6)	27.3 (2.1)	27.5 (2.1)	23.4 (1.8)
Mean figure of merit ^c		0.25; 0.49 after solvent flattening		
Refinement				
Resolution (Å)	50–1.7			
Reflections	36,301			
$R_{\text{cryst}}/R_{\text{free}}$ (%)	17.6/21.6			
Rms Δ bonds ^e	0.016 Å			
Rms Δ angles ^e	1.462°			
Average B factor (Å ²)	32.4			
Main chain dihedral angles				
Most favored	90.9%			
Allowed	9.1%			

^a Parentheses denote values for the highest-resolution shell.
^b $R_{\text{merge}} = \sum |I - \langle I \rangle| / \sum I$; I , intensity.
^c Mean figure of merit = $\langle |\sum_{\alpha} P(\alpha) e^{i\alpha} / \sum_{\alpha} P(\alpha)| \rangle$; α , phase; $P(\alpha)$, phase probability distribution.
^d $R_{\text{cryst}} = \sum |F_o - F_c| / \sum F_o$; F_o , observed structure factor amplitude; F_c , calculated structure factor amplitude.
^e Root mean square deviations from ideal values.

PTP functions and high substrate specificity (Alonso et al., 2004). The isolated catalytic domains of PTPs are generally active and often display rather broad substrate specificity. Regulatory domains often mediate specificity, regulation, and localization of PTPs. For example, the catalytic domain of PTP1B dephosphorylates substrates relatively indiscriminately when expressed in mammalian cells, but the full-length phosphatase containing an endoplasmic reticulum-targeting sequence dephosphorylates only a subset of these phosphoproteins (Barford et al., 1998). Due to the constitutively active nature of the catalytic domains, inhibition is the predominant regulatory mechanism associated with PTPs. A variety of inhibitory mechanisms have been described. Reversible oxidation of the catalytic cysteine is emerging as a general theme in PTP regulation and is well characterized for PTP1B (Salmeen et al., 2003; van Montfort et al., 2003). In the absence of phosphotyrosine substrate, the PTPs SHP-1 and SHP-2 are autoinhibited by their SH2 domains (Hof et al., 1998; Kuriyan and Cowburn, 1997; Yang et al., 2003). By contrast, the receptor protein tyrosine phosphatase (RPTP) RPTP α , and possibly other RPTPs, are inhibited by dimerization (Bilwes et al., 1996).

Here, we present the 1.7 Å resolution crystal structure of the 30 kDa tyrosine phosphatase PtpB of *Mtb*. Taken together, the structures of PtpA (Madhurantakam et al., 2005), PstP (Pullen et al., 2004), and PtpB provide a complete description of the so-called “eukaryotic-like” protein phosphatases in *Mtb*. PtpB contains the typical four-stranded parallel β sheet found in all PTP structures. Although PtpB displays the catalytic specificity for pTyr characteristic of conventional PTPs, the overall

fold is reminiscent of the dual-specificity phosphatases. These hybrid characteristics shared with other PTPs support the classification of PtpB as a member of a distinct PTP subfamily, which we propose to name the tyrosine-specific phosphatases (YSPs). The structure represents an inhibited conformation in which a structurally dynamic lid sequesters the active site from access to solvent. The lid contains a Phe residue that inserts into the active site pocket, and, together with the bound phosphate, mimics a phosphotyrosine substrate.

Results

PtpB Fold

Full-length PtpB was crystallized in the presence of the phosphate product, and the structure was determined using the multiwavelength anomalous diffraction method (Table 1). The structural model comprises residues 4–85, residues 117–275, a phosphate molecule bound in the catalytic site, and 265 water molecules. No electron density was visible for residues 86–116 or the last residue, Gly276. PtpB is a single-domain, globular protein with a central, four-stranded parallel β sheet surrounded by α helices (Figure 1).

PtpB shares the canonical, central, four-stranded parallel β sheet with other members of the PTP family. In comparison with conventional PTPs such as PTP1B or YopH, however, PtpB is deeply diverged and simplified structurally. In particular, PtpB lacks the antiparallel β strands that flank the central β sheet in PTP1B and YopH (Barford et al., 1994; Stuckey et al., 1994). Instead, PtpB contains several additional α helices. The closest structural homolog in the Protein Data Bank identified

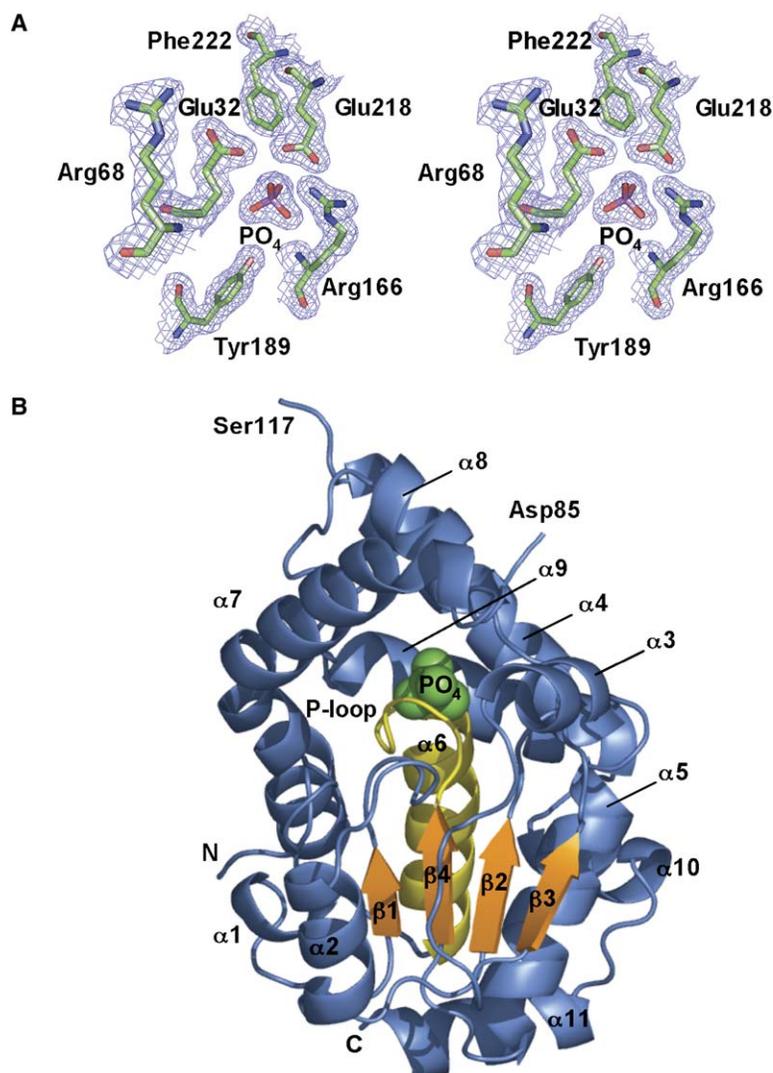


Figure 1. PtpB:PO₄ Overall Structure

(A) Stereo view of the bound phosphate and surrounding residues superimposed on the 1.7 Å resolution, $2F_o - F_c$ electron density map contoured at 1σ . The PO₄ is positioned adjacent to the Cys160 nucleophile (not shown) and is bound by the conserved Arg166 in the P loop. Phe222 from the lid is inserted into the active site cleft.

(B) Ribbon diagram of PtpB showing the central parallel β sheet (orange) and the buried active site. The P loop and the following helix, $\alpha 6$, are colored yellow. The P loop provides the catalytic cysteine nucleophile, as well as contacts that position and neutralize the phosphate (green). The lid (top) covers the active site, blocking the access of substrates.

by using the DALI server (Z score = 12) is the dual-specificity phosphatase At1g05000 from *Arabidopsis thaliana*. The C α root mean square deviation (rmsd) is 2.8 Å over 141 residues. This large rmsd is consistent with the low (20%) sequence identity of the plant and bacterial proteins (Figure 2A). The core β sheet, the active site P loop, and the following helix occur universally in the PTP family, including the LMW PTPs such as *Mtb* PtpA (Figures 2B and 3). The order of β strands that form the central β sheet in PtpB is $\beta 1$, $\beta 4$, $\beta 2$, $\beta 3$, similar to that of DSPs and conventional PTPs, but unlike that of any known LMW phosphatase (Figure 2B). In comparison to YopH and the *Arabidopsis* DSP, the general topology of PtpB shows insertions of helices as well as deletions of helices and β strands (Figures 2B and 3).

Sequence comparisons with other prokaryotic PTPs reveal two large insertions unique to the mycobacterial PtpBs. The first inserted region (85–117) forms a 31 residue disordered segment in the loop connecting $\beta 3$ and $\alpha 4$ (Figure 1B). This loop immediately follows the mobile “WPD” loop (FPD in PtpB) that harbors the conserved Asp that functions as a general acid in the PTPs. The beginning of the disordered loop is highly acidic, with five

aspartates in six residues. The end of the loop is rich in serine and glycine, which also are expected to enhance flexibility. Residues 95–101 are predicted to adopt a helical conformation (nnpredict).

The second insertion in PtpB relative to most homologs comprises residues 211–224, which form a large, two-helix lid that completely covers the active site (Figures 1B and 4). This lid must move out of the way to allow access to the active site, which is completely buried in the PtpB structure. The closest contacts to neighboring molecules in the lattice are more than 7 Å away from the lid, suggesting that the crystal packing does not determine the conformation of this segment.

The lid starts in the prominent and unusual $\alpha 7$, which forms a long, twisted helix that extends across one side of the protein (Figure 4B). Helix $\alpha 7$ contains two full turns of 3_{10} helix (190–196) preceding Pro197, which introduces a central bend between the 3_{10} segment and the C-terminal half of the helix. Not only is the 3_{10} helix embedded within $\alpha 7$, rather than occurring at a more typical location at one end, but the buried interface of the 3_{10} segment also is relatively polar. The packing surface of the N-terminal half of $\alpha 7$ contains Asp188,

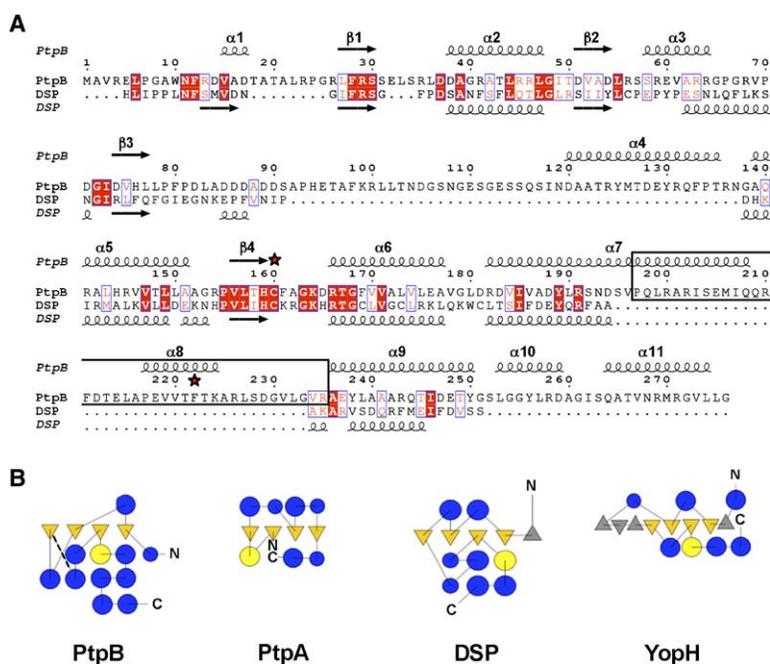


Figure 2. PtpB Adopts a Diverged PTP Fold
(A) Structure-sequence alignment of PtpB and the At1g05000 DSP from *Arabidopsis thaliana* (1XR1). The plant DSP is the most similar structure in the Protein Data Bank. BLAST detected no significant sequence similarity. The P loop and the following helix were used to align the structures. β strands (arrows) and α helices (spirals) are indicated adjacent to the sequences. The PtpB lid sequence is boxed. Identical residues are highlighted in red. Stars mark the catalytic cysteine and Phe222.

(B) Topology diagrams of PtpB, the *Mtb* LMW PtpA (1U2P), the *Arabidopsis* DSP (1XR1), and the *Yersinia* YopH conventional PTP (1QZ0). The core β sheet is colored orange, and the α helix following the P loop is colored yellow. The dashed line in the PtpB diagram indicates the disordered β 3- α 4 loop. The strand-P loop-helix motif characteristic of the active site occurs at the beginning of the LMW PTP sequences, but in the latter half of the conventional PTPs and DSPs. The distinct PtpB topology shares features with the conventional PTPs and DSPs.

Tyr189, Ser192, and Ser195. Asp188 forms a buried ion pair with Arg29. The buried Ser192 forms buried hydrogen bonds with Asn11, Arg188, and the backbone of the P loop (Gly163 O). This P loop carbonyl also contacts the buried guanidinium group of Arg29. Arg29 occurs in one of two unusual stacks of three adjacent arginines in the PtpB structure, involving Arg191, Arg13, and Arg29 (Figure 4C). The other arginine stack, containing residues Arg68, Arg35, and Arg64, extends under the lid.

A seven residue loop (211–217), starting with the unusually exposed Phe211, links α 7 to α 8 (Figure 4), and α 8 is joined to α 9 by a meandering, 11 residue connector. Helix 8 contains Phe222, projecting into the active site. Superposition of the structures of the PtpB:PO₄ complex and the YopH complex with a phosphotyrosyl mimetic-containing peptide (Phan et al., 2003) revealed that access to the active site is blocked by Phe222 in a manner that is reminiscent of phosphotyrosine binding (Figure 3). The C4 of Phe222 is nearly in van der Waals contact with the bound PO₄ (C4-O distance = 4.3 Å).

Active Site

In contrast to the diverged fold, PtpB displays the same active site configuration as other PTPs. The catalytic cysteine, Cys160, is positioned within the P loop consensus sequence HisCysXAlaGlyXXArg between β 5 and α 6. The Cys160 thiol group is positioned for nucleophilic attack on the phosphate (S-P distance = 3.52 Å), which makes characteristic hydrogen bonded ion pairs with the buried side chains of Lys164 and Arg166. The phosphate bound in the active site is positioned at the N terminus of α 6, aligned with the axis of the helix. The dipole of α 6, the partial positive charges of the P loop amide hydrogens, and Arg166 all contribute to a binding pocket complementary to the phosphate. The Cys160 carbonyl oxygen contacts the amide nitrogen of Glu32, the side chain of which forms buried ion pairs with Arg56 and Arg64.

Asp82, corresponding to the general acid Asp181 in PTP1B, is positioned within a modified WPD motif with the sequence FPD. The FPD motif is located in a loop connecting β 3 and α 4. This FPD variation also occurs in the mycobacterial homologs of PtpB. In comparison with structures of PTP1B with the WPD loop in closed (1PTY) and open conformations (1NO6), the FPD loop of PtpB approaches a closed conformation that brings it into the active site in a position for catalysis.

The active site region comprises nearly all of the segments conserved in a sample of 60 PtpB homologs (Figure 5A). The most conserved elements include the P loop, the first turn of the following helix, and the other loops (including the FPD loop) that ring the phosphate binding site. Helix α 3, which is expected to influence the depth of the substrate binding pocket when the lid opens, also shows conservation. Other conserved elements include the sequence just preceding the kink in the long helix 7 (Figure 5A). The buried residues of this sequence contact the P loop, positioning it within the active site and presumably influencing the electrostatic potential at the catalytic center. It is striking that the exposed surface of PtpB shows generally low conservation in homologs (Figure 5B), and this pattern supports the idea that the conformation undergoes a large change to expose conserved functional sites.

Regulation

The activity of PtpB was assessed in solution and in the crystals. Addition of the chromogenic phosphotyrosine mimic, p-nitrophenylphosphate (pNPP), to PtpB in solution resulted in the production of bright yellow color. The K_M of PtpB for pNPP is ~10 mM at pH 7.5 (data not shown), similar to the K_M for pNPP reported for other PTPs (Sun et al., 2003; Villa et al., 2005). The pH optima of PtpB with pNPP or the phosphopeptide RRLIEDAEPYAARG were found to be near pH 5.5 (data not shown), in agreement with the pH optimum reported previously

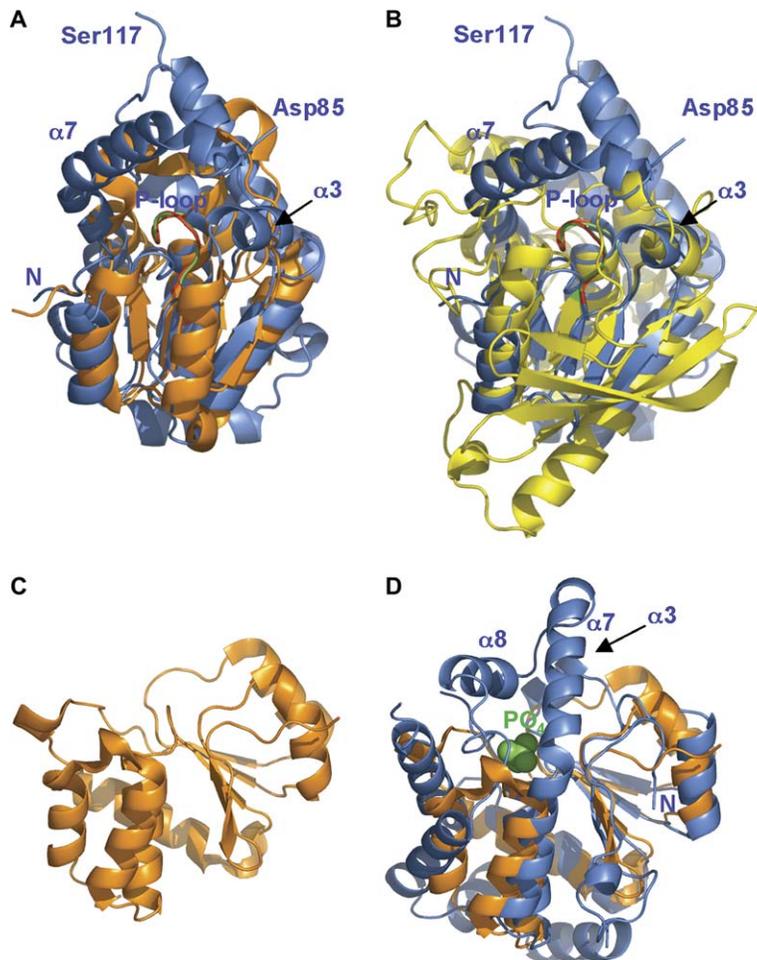


Figure 3. Ribbon Diagrams of Structural Superpositions Show the Distinct Fold of PtpB

Structures were superimposed by using the P loop and the following helix.

(A) PtpB (1YWF, blue) and the *Arabidopsis* DSP (1XRI, orange). PtpB $\alpha 3$ helps frame a deeper active site pocket.

(B) PtpB and YopH (1QZ0, yellow) in an orientation similar to (A). YopH contains a larger β sheet and additional helices. The PtpB lid and the disordered loop flanked by Asp85 and Ser117 have no sequence or structural analogs in the conventional PTPs.

(C) The *Arabidopsis* DSP (1XRI) contains a characteristically shallow, accessible active site (top).

(D) Superposition of the DSP with PtpB showing a blockade of the PtpB active site by the lid. The phosphate (green) binds to the P loop. The lid sequences and $\alpha 3$ (back) extend above the active site.

for PtpB-catalyzed dephosphorylation of myelin basic protein phosphorylated on Tyr by Src kinase (Koul et al., 2000).

While PtpB catalyzed pNPP hydrolysis in the pH 5.5 mother liquor used for crystallization, addition of pNPP to PtpB crystals resulted in no color change, indicating that the crystal structure represents a catalytically inactive form of the enzyme. The lack of catalytic activity in the crystals is consistent with the striking inaccessibility of the substrate binding site, which is covered by the lid (Figure 1). These results suggest that the lid forms a flexible element in solution that regulates access of substrates and small molecules to the active site.

Due to the reactivity of the Cys nucleophile, PTPs are generally sensitive to regulation or inactivation *in vivo* by reactive oxygen species (ROS [Salmeen et al., 2003; Barford, 2004]). PtpB is likely to encounter ROS in host macrophages, so we probed the rate of oxidative inactivation of PtpB by hydrogen peroxide (Figure 6). *Mtb* PtpA, a LMW PTP with an open active site (Madhuran-takam et al., 2005), was used as a positive control for oxidative inactivation. Treatment with 62.5 μM H_2O_2 for 20 min reduced PtpA activity to 73%, which indicates ROS sensitivity comparable to that of the mouse RPTPa-D1 domain (Groen et al., 2005). In contrast, 93% of PtpB activity remained after similar treatment. Exposure to 250 μM H_2O_2 for 20 min inactivated over 80% of the

PtpA, but only 37% of the PtpB (Figure 6). Thus, in comparison to PtpA, PtpB was resistant to oxidative inactivation.

Discussion

Despite the absence of a recognizable tyrosine kinase, the genome of *Mtb* encodes two tyrosine phosphatases, PtpA and PtpB. Sequence comparisons reveal the presence of PtpB orthologs in *M. bovis*, *M. microti*, *M. ulcerans*, *M. marinum*, and *M. avium*. No PtpB ortholog is present in *M. smegmatis* or *M. leprae*. The closest PtpB homolog (34% sequence identity) outside mycobacteria was found in the Actinomycete *Nocardia farci-ana*. The lid and the long $\alpha 4$ - $\beta 4$ loop sequences, in particular, are restricted to the mycobacterial PtpBs. Thus, true PtpB orthologs appear to be restricted to pathogenic mycobacteria. Although the *Mtb* PtpB sequence is readily recognized as a PTP, over 100 of the closest homologs occur in bacteria rather than eukaryotes. Because the sequence shows no extended homology outside the active site to structures in the PDB, the PtpB structure defines a distinct, widespread solution to the problem of protein tyrosine dephosphorylation function.

Despite the apparent specificity for pTyr substrates (Koul et al., 2000), PtpB is structurally diverged from the conventional PTPs. None of the over two dozen

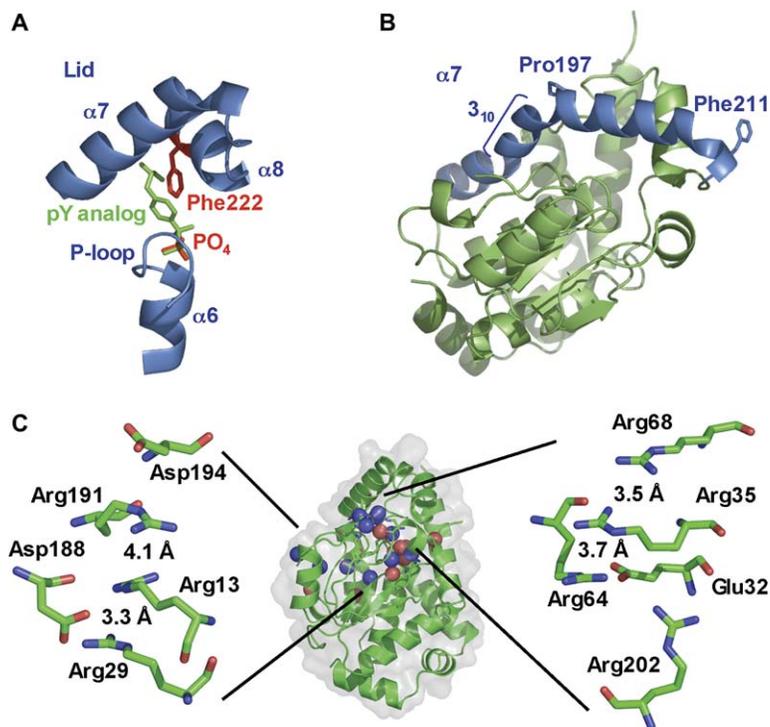


Figure 4. A Structurally Flexible Lid Covers the PtpB Active Site, Sequestering the Product Phosphate from Solvent

The PtpB structure displays five potential mechanisms of promoting conformational switching: substrate mimicry, an extended 3_{10} helix, exposed hydrophobic groups, buried polar interfaces, and buried charges.

(A) Substrate mimicry suggested by the superposition of PtpB with the substrate analog from a YopH:peptide complex (1QZ0, green). The combination of PtpB Phe222 and PO_4 (both highlighted in red) structurally resembles the phosphotyrosyl-substrate mimetic in the YopH complex superimposed with the P loop and the following helix. Closer approach of Phe222 and the phosphate is limited by their van der Waals radii. The PtpB lid buries the bound phosphate.

(B) Structural features in PtpB (green) associated with the putative metastable, closed conformation of the lid. Helix $\alpha 7$ (blue) contains two turns of 3_{10} helix, a proline bend, exposed hydrophobic residues, and a buried polar surface (not shown).

(C) Arginine stacks and buried charged residues associated with the lid: Asp54, Asp82, Asp188, Glu32, Lys164, and Arg13, Arg29, Arg56, Arg64, Arg202 (center: Asp OD2 and Glu OE2 atoms, red spheres; Arg NH2 and Lys NZ atoms, blue spheres). The two stacks of arginines are shown reoriented for clarity, with the distances between the Arg residues indicated. The buried, charged side chains in these clusters include Asp188, Arg13, and Arg29 (left) and Arg203 and Glu32 (right).

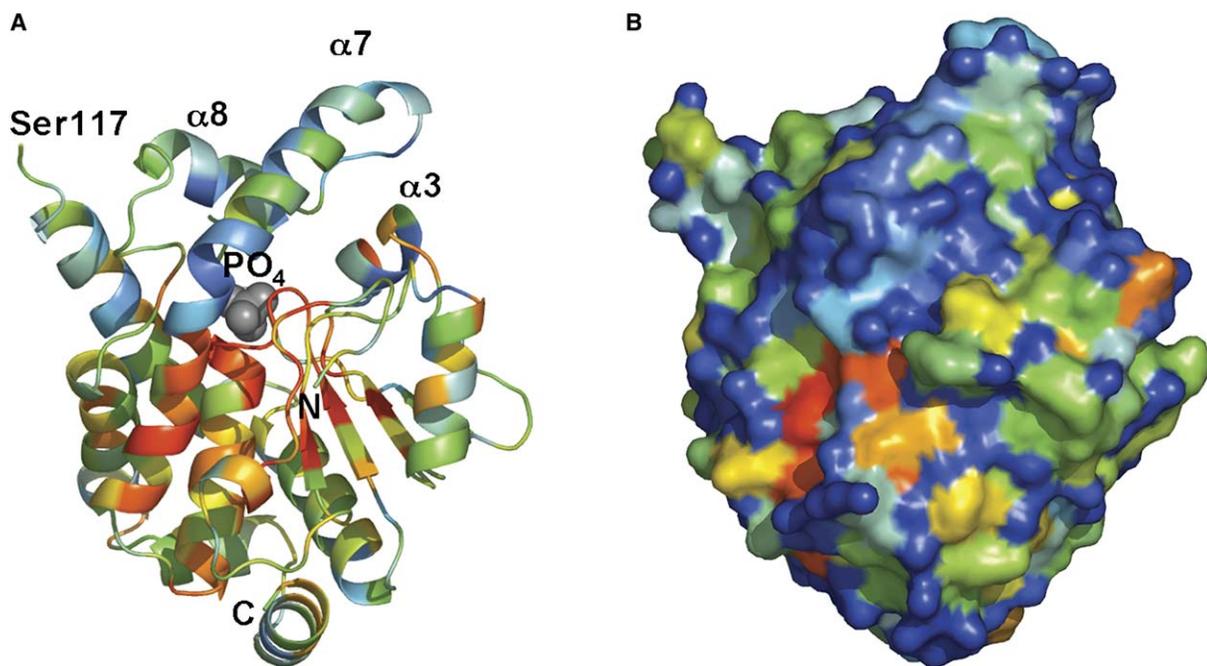


Figure 5. Sequence Conservation Highlights Functional Sites in PtpB

(A and B) Conservation in 60 representative homologs identified using a BLAST search (E value < 0.001) of 356 microbial genomes. (A) Ribbon diagram and (B) surface view. Relative to Figure 1B, the view corresponds to an $\sim 45^\circ$ right rotation around the vertical axis. Red corresponds to high conservation, and blue corresponds to low conservation. Gradations through the visual spectrum represent intermediate values. The most conserved residues surround the position of the bound phosphate (green) and the P loop.

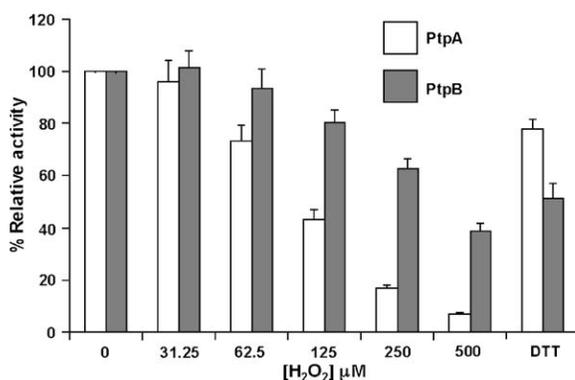


Figure 6. PtpB Resists Oxidation by H₂O₂

PtpA and PtpB were assayed using photometric detection of pNPP hydrolysis after a 20 min treatment with the indicated concentration of H₂O₂. Error bars indicate the standard deviation of three measurements. PtpA was more susceptible to oxidation. The addition of 2 mM DTT partly restored activity, indicating the presence of the singly oxidized species.

PTP structures in the PDB nor the structure of the PTEN phosphatidyl-inositol-3-phosphate phosphatase show close structural homology to PtpB. An evolutionary tree (Edgar and Sjolander, 2003) constructed using the sequences of selected PTPs in the PDB and homologs identified with BLAST suggests that PtpB may represent a distinct branch that shares features of conventional PTPs as well as DSPs (Figure S1; see the Supplemental Data available with this article online).

The structure suggests that PtpB can be thought of as a simplified conventional PTP or as a DSP with limited substrate specificity. Compared to the conventional PTPs exemplified by PTP1B and YopH, PtpB contains a smaller β sheet, as well as deletions and insertions of helices and loops that are more reminiscent of the DSPs (Figure 2B). Compared to the *Arabidopsis* DSP At1g05000, PtpB contains similar core secondary structural elements (Figures 3A and 3D). A principal difference is that α 3 and the following loop in PtpB coincide structurally with a loop and α 2 in the plant phosphatase (Figures 2A and 3A). PtpB also contains the inserted sequences that form the α 4- β 4 loop and the lid covering the active site, in contrast to the open, shallow substrate binding site in the DSP (Figures 3C and 3D). The large backbone rmsd of PtpB compared to any available PTP structure and the evolutionary tree constructed from PTP sequences are consistent with the idea that PtpB represents a separate PTP subfamily. We propose to call this subfamily the YSPs, for tyrosine-specific phosphatases. This name reflects the structural similarities to DSPs and the catalytic specificity for phosphotyrosine protein substrates.

The distinct substrate specificities of PTP subfamilies are thought to arise from the depth of the active site pocket (Barford et al., 1998). The pTyr specificity of the conventional PTPs is associated with a deep pocket that excludes the smaller pSer and pThr residues. In contrast, the DSPs have a more open active site that allows access of pSer, pThr, and pTyr to the catalytic Cys residue (Figure 3C). PtpB clearly shows exclusive specificity for phosphotyrosine substrates in vitro (Koul et al., 2000). Several loops ringing the PtpB active site (resi-

dues 9–11, 30–33, 81–83) show conformations similar to those seen in conventional phosphatases and DSPs (Figure 3). In contrast, PtpB α 3, which coincides structurally with an extended loop in the *Arabidopsis* DSP, helps frame a deeper pocket similar to conventional PTPs. The conventional PTPs also create a deeper active site due to loops that coincide structurally, but not topologically, to the PtpB lid. Thus, the position of the lid in the open conformation also may contribute to the pTyr specificity.

The closed conformation of the lid places the edge of Phe222, which protrudes into the active site from α 8, near the bound PO₄. This arrangement resembles the phosphotyrosine substrate position expected from the superposition of PtpB with PTP structures with substrates or inhibitors bound (Figure 4A). This intramolecular substrate mimicry inhibits PtpB activity in the crystals by blocking access to the active site.

The blockade of the active site by the lid (Figures 1B, 3, and 4A) indicates that an extensive structural change is required for substrate binding. This conformational change is facile in solution. The K_M of PtpB for pNPP (~10 mM at pH 7.5), for example, was similar to those of other PTPs, such as YopH or PTPL1 (1.7 mM and 17.9 mM, respectively [Sun et al., 2003; Villa et al., 2005]). Compared to other PTPs (Salmeen et al., 2003; Groen et al., 2005), however, PtpB was significantly more resistant to inactivation by hydrogen peroxide (Figure 6). While 50% inactivation of PtpB at pH 7.5 required a 20 min incubation with 500 μ M H₂O₂, PTP1B was inactivated to a similar extent after 10 min at 160 μ M H₂O₂ (Salmeen et al., 2003). After a 20 min treatment, 50% inactivation of RPTPa-D1 was reported at 31–62 μ M H₂O₂, and >50% inactivation of the RPTPa-C2 domain required <31 μ M H₂O₂ (Groen et al., 2005). Although each of these measurements was done at pH 7.5, the rates of oxidative inactivation in vivo will be influenced by the pH (Groen et al., 2005). Despite the caveat that such data collected in different laboratories may not provide a firm basis for quantitative comparisons, PtpB qualitatively shows the highest resistance to H₂O₂ inactivation of any PTP examined to date (Figure 6).

This resistance to oxidative inactivation supports the hypothesis that the PtpB lid may provide a dynamic filter that excludes ROS. If closure of the lid is rapid compared to oxidative inactivation, for example, H₂O₂ may be effectively excluded from the active site, preserving PtpB activity in infected cells. In contrast, bound protein substrates or modifications of PtpB that stabilize the open conformation of the lid would promote substrate turnover. Thus, the occurrence of the autoinhibited conformation in the crystals raises the possibility that lid opening and closing may be regulated in vivo.

Several features of the PtpB structure are consistent with the idea that the lid conformation is dynamic. The two turns of the 3_{10} helix embedded within α 7 (Figures 1B and 4B) constitute an unusual, strained conformation (Rohl and Doig, 1996). This 3_{10} -helical segment might serve as an overwound spring that supports rearrangement of the neighboring helices upon unwinding. Consistent with this idea, the 3_{10} helix sequence contains a two residue deletion that is specific to the mycobacterial orthologs containing the lid sequence. Moreover, α 7 and the lid contain buried polar contacts, including two

unusual stacks of three or four Arg residues and two buried carboxylates (Figure 4C). Helix $\alpha 7$ also displays exposed hydrophobic residues such as Val184, Ala187, Leu190, Ala 201, and Phe211. Opening the lid may eliminate these unusual features through extensive rearrangements of helices 7–9 and perhaps new hydrophobic contacts with target proteins.

Due to the high similarity of the active site in PTPs, conventional PTPs and DSPs typically contain additional domains or localization signals that confer substrate specificity (Barford et al., 1998). In the PtpB structure, the additional segments include the lid module and the disordered, 31 residue $\beta 3$ - $\alpha 4$ loop. The $\beta 3$ - $\alpha 4$ loop protrudes from a structure adjacent to the lid into a position that allows it to move toward the active site when the lid moves away. These two diverged segments are positioned to contribute to substrate binding and/or regulation.

The PtpB fold is much more distinct from its eukaryotic counterparts than the structures of other *Mtb* phosphosignaling proteins, including the Ser/Thr kinase PknB (Young et al., 2003; Ortiz-Lombardia et al., 2003), the Ser/Thr phosphatase PstP (Pullen et al., 2004), and the LMW PtpA (Madhurantakam et al., 2005). Indeed, the entire PTP family shows more elaboration of secondary structure than many other protein families (Figures 2B and 3). This divergence may arise from selection for changes in substrate recognition and/or regulation. The unusual, extensive structural differences in the DSPs and conventional PTPs implies that variation in the number and length of secondary structural elements is disfavored in most other protein families.

In contrast to the diverged fold, the active site closely resembles that of other PTPs. These sequence and structural similarities support the idea of a common catalytic mechanism. Consistent with this conclusion, mutations of the presumptive nucleophile, Cys160, and the general acid, Asp82, greatly reduced catalytic activity (data not shown). These results suggest that the Cys160Ser and Asp82Ala mutants might be used to trap complexes with *in vivo* substrates, in a manner analogous to other PTPs (Flint et al., 1997).

Central questions about PtpB concern the identity of the substrate phosphoproteins in host cells and the mechanism by which the enzyme promotes virulence. Tyrosine phosphorylation plays a key role in many innate and adaptive immune functions in macrophages, providing many possible targets for PtpB to facilitate immune escape and *Mtb* survival. The unique features of the PtpB structure in comparison to the virulence factors YopH and SptP are consistent with the model that PtpB targets different substrates. Not only does PtpB act later in infection than the *Yersinia* and *Salmonella* PTPs (Singh et al., 2003), but, the PtpB domain structure is also distinct. YopH and SptP contain a C-terminal phosphatase domain and an N-terminal domain that mediates targeting (YopH) or GTPase-activating protein (GAP) activity (SptP). *Yersinia* also encodes a GAP, YopE, which colocalizes with YopH in infected cells and acts synergistically with YopH (Cornelis, 2002). In contrast, PtpB contains a simpler PTP fold with no appended domains. The PtpB structure suggests that rearrangements of the lid and/or the disordered $\beta 3$ - $\alpha 4$ loop adjacent to the active site are required for substrate

binding and targeting. The unique burial of the active site may protect PtpB from oxidative inactivation or provide a mechanism to regulate substrate binding. The structure provides a framework to test these ideas by mutagenesis. The PtpB active site architecture also enables the design of substrate-trapping mutants to help identify the host targets (Flint et al., 1997) and lays the groundwork to develop specific inhibitors.

Experimental Procedures

Cloning, Protein Expression, and Purification

The full-length PtpB gene was amplified from genomic DNA of the *Mtb* H37Rv strain (Colorado State University). PCR products were cloned in frame with an N-terminal His₆ tag into the pET28b expression vector (Novagen). PtpB mutants were generated using the Quik-change protocol (Stratagene), and all constructs were confirmed by DNA sequencing. For protein expression, plasmids were transformed into BL21 (DE3)-CodonPlus cells (Stratagene) and grown in Terrific Broth containing 50 μ g/ml kanamycin and 34 μ g/ml chloramphenicol at 37°C to an OD₆₀₀ of 0.5. After adding of 20 μ M IPTG, the cultures were grown for an additional 20 hr at 20°C. The cells were harvested by centrifugation and resuspended in 20 mM NaH₂PO₄ (pH 7.5), 100 mM NaCl, and 5% glycerol. The cells were lysed by sonication on ice and centrifuged at 15,000 \times g for 1 hr. The protein was purified by immobilized metal affinity chromatography, size exclusion chromatography (Superdex S75), and ion exchange chromatography (MonoQ). The protein eluted from the ion exchange column in 20 mM Tris (pH 7.5), 150 mM NaCl and was concentrated to 5 mg/ml. SeMet PtpB was expressed as described (Van Duyne et al., 1993) and was purified as described above for PtpB.

Crystallization

Crystals were grown by the vapor diffusion method in hanging drops at 4°C. A total of 1 μ l PtpB (5 mg/ml) was added to 1 μ l 0.05 M KH₂PO₄, 20% PEG 8000, 10 mM sodium acetate (pH 5.5). For data collection, crystals were immersed in 0.05 M KH₂PO₄, 20% PEG 8000, 10 mM sodium acetate, and 25% glycerol, mounted on a loop, and flash frozen in liquid nitrogen.

Structure Determination

Diffraction data were collected at the Lawrence Berkeley National Laboratory Advanced Light Source Beamline 8.3.1. Native data were collected to 1.7 Å resolution at 100° K by using 1.11587 Å X-rays. Phases were obtained using MAD data collected from an SeMet crystal at the Se peak, high-remote, and low-remote energies. The crystals had the symmetry of space group I4₁, with unit cell dimensions a = 113.068 Å, b = 113.068 Å, c = 53.292 Å, and one protein molecule in the asymmetric unit. Data were reduced by using HKL2000 (Otwinowski and Minor, 1997). Solve and Resolve (Terwilliger, 2003) were used to identify three of the four SeMet sites and calculate initial phases. The initial model was built with ARP/wARP (Morris et al., 2002). The final model was obtained through repeated rebuilding with O (Jones et al., 1991) and refinement with REFMAC (Murshudov et al., 1997). The R_{free} was calculated by using a random 5% of the data.

Phosphatase Activity Assays

Activity with the substrate p-nitrophenyl phosphate (pNPP) for K_M determination was measured photometrically at A₄₀₅. Absorbance was measured after incubation of 1 μ M enzyme with 25 mM pNPP in 20 mM Tris, 150 mM NaCl (pH 7.5) for various times. Steady-state kinetic parameters were calculated by using SigmaPlot (Systat Software, Inc.). Activity of PtpB with the phosphopeptide RRLIEDAE-pYAARG (Upstate) was determined by using the malachite green assay (Baykov et al., 1988). Equal volumes of 30 mM ammonium molybdate in 4 M HCl and 0.05% (w/v) malachite green were incubated for 30 min at room temperature (RT), and 0.01% Tween-20 was added. The phosphatase reactions were stopped after 30 min by adding a 2-fold excess of malachite green solution and were read at A₆₅₀. Both assays were read by using a SPECTRAMax 190 spectrophotometer (Molecular Devices).

For oxidation measurements (Groen et al., 2005), PtpA and PtpB were bound to Ni-NTA Superflow resin (Qiagen) and fully reduced by incubating in buffer (20 mM Tris [pH 7.5], 100 mM NaCl) containing 10 mM DTT for 20 min. The resin was washed three times in buffer without DTT, and the protein was eluted with 125 mM imidazole. To remove imidazole and traces of DTT, the proteins were dialyzed against 10,000 volumes of buffer. The enzymes were incubated with different amounts of H₂O₂ for 20 min at room temperature. Activity was measured by adding 10 mM p-nitrophenylphosphate in 20 mM Tris (pH 7.5), 100 mM NaCl for 5 min at RT. The reaction was quenched by adding 0.4 M NaOH, and absorbance at A₄₄₀ was measured.

Structure Analysis

Sequences of homologs were identified with BLAST, and additional PTP sequences were retrieved from the PDB. Multiple sequence alignments were created with MUSCLE (Edgar, 2004) and edited manually. An evolutionary tree was calculated by using SATCHMO (Edgar and Sjolander, 2003). The results were used to color the solvent-accessible surface of PtpB by using ConSurf (Glaser et al., 2003). The electrostatic surface potential of PtpB was displayed by using Pymol. Topology diagrams were generated with TOPS (Michalopoulos et al., 2004) and edited manually.

Supplemental Data

Supplemental Data including an evolutionary tree of 51 eukaryotic and prokaryotic PTPs are available at <http://www.structure.org/cgi/content/full/13/11/1625/DC1/>.

Acknowledgments

We thank K. Pullen, P. Barth, S. Brenner, and Y. Av-Gay for helpful discussions and J. Holton for help with X-ray data collection. We are indebted to T. Terwilliger and the TB Structural Genomics Consortium for support. This work was supported by a grant from the National Institutes of Health. The Advanced Light Source Beamline 8.3.1 was funded by the National Science Foundation, the University of California, and Henry Wheeler. H.-L.N. was supported by a National Institutes of Health Post-doctoral Fellowship. C.G. gratefully acknowledges a postdoctoral fellowship from the New York Community Trust-Heiser Fund.

Received: May 9, 2005

Revised: July 18, 2005

Accepted: July 23, 2005

Published: November 8, 2005

References

- Alonso, A., Sasin, J., Bottini, N., Friedberg, I., Osterman, A., Godzik, A., Hunter, T., Dixon, J., and Mustelin, T. (2004). Protein tyrosine phosphatases in the human genome. *Cell* 117, 699–711.
- Barford, D. (2004). The role of cysteine residues as redox-sensitive regulatory switches. *Curr. Opin. Struct. Biol.* 14, 679–686.
- Barford, D., Flint, A.J., and Tonks, N.K. (1994). Crystal structure of human protein tyrosine phosphatase 1B. *Science* 263, 1397–1404.
- Barford, D., Das, A.K., and Egloff, M.P. (1998). The structure and mechanism of protein phosphatases: insights into catalysis and regulation. *Annu. Rev. Biophys. Biomol. Struct.* 27, 133–164.
- Baykov, A.A., Evtushenko, O.A., and Avaeva, S.M. (1988). A malachite green procedure for orthophosphate determination and its use in alkaline phosphatase-based enzyme immunoassay. *Anal. Biochem.* 171, 266–270.
- Bilwes, A.M., den Hertog, J., Hunter, T., and Noel, J.P. (1996). Structural basis for inhibition of receptor protein-tyrosine phosphatase- α by dimerization. *Nature* 382, 555–559.
- Black, D.S., and Bliska, J.B. (1997). Identification of p130Cas as a substrate of *Yersinia* YopH (Yop51), a bacterial protein tyrosine phosphatase that translocates into mammalian cells and targets focal adhesions. *EMBO J.* 16, 2730–2744.
- Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E., 3rd, et al.

(1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393, 537–544.

Cornelis, G.R. (2002). The *Yersinia* Ysc-Yop 'type III' weaponry. *Nat. Rev. Mol. Cell Biol.* 3, 742–752.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797.

Edgar, R.C., and Sjolander, K. (2003). SATCHMO: sequence alignment and tree construction using hidden Markov models. *Bioinformatics* 19, 1404–1411.

Flint, A.J., Tiganis, T., Barford, D., and Tonks, N.K. (1997). Development of 'substrate-trapping' mutants to identify physiological substrates of protein tyrosine phosphatases. *Proc. Natl. Acad. Sci. USA* 94, 1680–1685.

Fu, Y., and Galan, J.E. (1998). The *Salmonella typhimurium* tyrosine phosphatase SptP is translocated into host cells and disrupts the actin cytoskeleton. *Mol. Microbiol.* 27, 359–368.

Fu, Y., and Galan, J.E. (1999). A salmonella protein antagonizes Rac-1 and Cdc42 to mediate host-cell recovery after bacterial invasion. *Nature* 401, 293–297.

Glaser, F., Pupko, T., Paz, I., Bell, R.E., Bechor-Shental, D., Martz, E., and Ben-Tal, N. (2003). ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics* 19, 163–164.

Groen, A., Lemeer, S., van der Wijk, T., Overvoorde, J., Heck, A.J., Ostman, A., Barford, D., Slijper, M., and den Hertog, J. (2005). Differential oxidation of protein-tyrosine phosphatases. *J. Biol. Chem.* 280, 10298–10304.

Hof, P., Pluskey, S., Dhe-Paganon, S., Eck, M.J., and Shoelson, S.E. (1998). Crystal structure of the tyrosine phosphatase SHP-2. *Cell* 92, 441–450.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* 47 (Pt 2), 110–119.

Kaniga, K., Uralil, J., Bliska, J.B., and Galan, J.E. (1996). A secreted protein tyrosine phosphatase with modular effector domains in the bacterial pathogen *Salmonella typhimurium*. *Mol. Microbiol.* 21, 633–641.

Koul, A., Choidas, A., Treder, M., Tyagi, A.K., Drica, K., Singh, Y., and Ullrich, A. (2000). Cloning and characterization of secretory tyrosine phosphatases of *Mycobacterium tuberculosis*. *J. Bacteriol.* 182, 5425–5432.

Kuriyan, J., and Cowburn, D. (1997). Modular peptide recognition domains in eukaryotic signaling. *Annu. Rev. Biophys. Biomol. Struct.* 26, 259–288.

Madhurantakam, C., Rajakumara, E., Mazumdar, P.A., Saha, B., Mitra, D., Wiker, H.G., Sankaranarayanan, R., and Das, A.K. (2005). Crystal structure of low-molecular-weight protein tyrosine phosphatase from *Mycobacterium tuberculosis* at 1.9-Å resolution. *J. Bacteriol.* 187, 2175–2181.

Michalopoulos, I., Torrance, G.M., Gilbert, D.R., and Westhead, D.R. (2004). TOPS: an enhanced database of protein structural topology. *Nucleic Acids Res.* 32, D251–D254.

Morris, R.J., Perrakis, A., and Lamzin, V.S. (2002). ARP/wARP's model-building algorithms. I. The main chain. *Acta Crystallogr. D Biol. Crystallogr.* 58, 968–975.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* 53, 240–255.

Ortiz-Lombardia, M., Pompeo, F., Boitel, B., and Alzari, P.M. (2003). Crystal structure of the catalytic domain of the PknB serine/threonine kinase from *Mycobacterium tuberculosis*. *J. Biol. Chem.* 278, 13094–13100.

Otwinowski, Z., and Minor, W. (1997). Processing of x-ray diffraction data collected in oscillation mode. In *Methods in Enzymology*, Volume 276, C.W. Carter, Jr., and R. Sweet, eds. (New York: Academic Press), pp. 307–326.

Phan, J., Lee, K., Cherry, S., Tropea, J.E., Burke, T.R., Jr., and Waugh, D.S. (2003). High-resolution structure of the *Yersinia pestis* protein tyrosine phosphatase YopH in complex with a

phosphotyrosyl mimetic-containing hexapeptide. *Biochemistry* **42**, 13113–13121.

Pullen, K.E., Ng, H.L., Sung, P.Y., Good, M.C., Smith, S.M., and Alber, T. (2004). An alternate conformation and a third metal in PstP/Ppp, the *M. tuberculosis* PP2C-family Ser/Thr protein phosphatase. *Structure (Camb)* **12**, 1947–1954.

Rohl, C.A., and Doig, A.J. (1996). Models for the 3(10)-helix/coil, p-helix/coil, and α -helix/3(10)-helix/coil transitions in isolated peptides. *Protein Sci.* **5**, 1687–1696.

Salmeen, A., Andersen, J.N., Myers, M.P., Meng, T.C., Hinks, J.A., Tonks, N.K., and Barford, D. (2003). Redox regulation of protein tyrosine phosphatase 1B involves a sulphenyl-amide intermediate. *Nature* **423**, 769–773.

Singh, R., Rao, V., Shakila, H., Gupta, R., Khera, A., Dhar, N., Singh, A., Koul, A., Singh, Y., Naseema, M., et al. (2003). Disruption of mptpB impairs the ability of *Mycobacterium tuberculosis* to survive in guinea pigs. *Mol. Microbiol.* **50**, 751–762.

Stuckey, J.A., Schubert, H.L., Fauman, E.B., Zhang, Z.Y., Dixon, J.E., and Saper, M.A. (1994). Crystal structure of *Yersinia* protein tyrosine phosphatase at 2.5 Å and the complex with tungstate. *Nature* **370**, 571–575.

Sun, J.P., Wu, L., Fedorov, A.A., Almo, S.C., and Zhang, Z.Y. (2003). Crystal structure of the *Yersinia* protein-tyrosine phosphatase YopH complexed with a specific small molecule inhibitor. *J. Biol. Chem.* **278**, 33392–33399.

Terwilliger, T.C. (2003). SOLVE and RESOLVE: automated structure solution and density modification. *Methods Enzymol.* **374**, 22–37.

Van Duyn, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L., and Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.* **229**, 105–124.

van Montfort, R.L., Congreve, M., Tisi, D., Carr, R., and Jhoti, H. (2003). Oxidation state of the active-site cysteine in protein tyrosine phosphatase 1B. *Nature* **423**, 773–777.

Vergne, I., Chua, J., Lee, H.H., Lucas, M., Belisle, J., and Deretic, V. (2005). Mechanism of phagolysosome biogenesis block by viable *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. USA* **102**, 4033–4038.

Villa, F., Deak, M., Bloomberg, G.B., Alessi, D.R., and van Aalten, D.M. (2005). Crystal structure of the PTPL1/FAP-1 human tyrosine phosphatase mutated in colorectal cancer: evidence for a second phosphotyrosine substrate recognition pocket. *J. Biol. Chem.* **280**, 8180–8187.

Walburger, A., Koul, A., Ferrari, G., Nguyen, L., Prescianotto-Baschong, C., Huygen, K., Klebl, B., Thompson, C., Bacher, G., and Pieters, J. (2004). Protein kinase G from pathogenic mycobacteria promotes survival within macrophages. *Science* **304**, 1800–1804.

Yang, J., Liu, L., He, D., Song, X., Liang, X., Zhao, Z.J., and Zhou, G.W. (2003). Crystal structure of human protein-tyrosine phosphatase SHP-1. *J. Biol. Chem.* **278**, 6516–6520.

Yao, T., Meccas, J., Healy, J.I., Falkow, S., and Chien, Y. (1999). Suppression of T and B lymphocyte activation by a *Yersinia pseudotuberculosis* virulence factor, YopH. *J. Exp. Med.* **190**, 1343–1350.

Young, T.A., Delagoutte, B., Endrizzi, J.A., Falick, A.M., and Alber, T. (2003). Structure of *Mycobacterium tuberculosis* PknB supports a universal activation mechanism for Ser/Thr protein kinases. *Nat. Struct. Biol.* **10**, 168–174.

Accession Numbers

The coordinates and structure factors were deposited in the Protein Data Bank (accession number 1YWF).